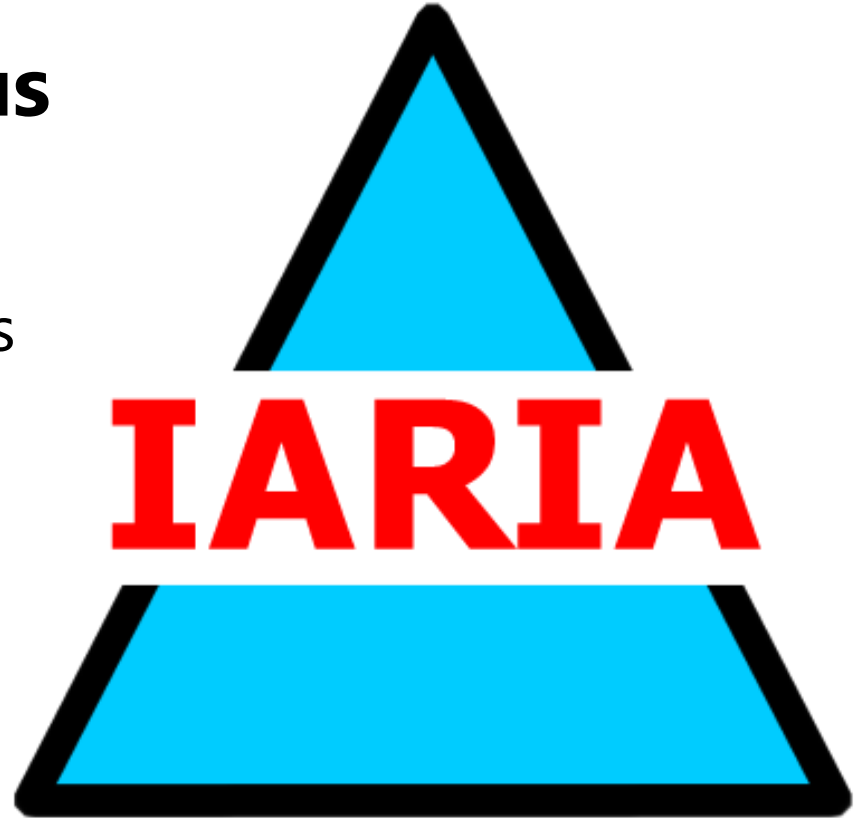
# AI-Assisted UML Modeling for Serious Mental Illness Crisis Management—

## Balancing Automation and Human Oversight: A Comparative Analysis of UML Diagramming Methods


Authors: Kerry Gilder, John Chelsom
Presented by: Kerry Gilder

**Applied Health Informatics Program**
**Fordham University**
**New York, USA**

kgilder@fordham.edu
jchelsom@fordham.edu

# Author Bios

## Kerry Gilder, M.S. (CASAC-T)

Graduate Student, Applied Health Informatics, Fordham University

Kerry is a multidisciplinary professional based in the New York Metro area with a background in business, cybersecurity, and behavioral health. She is a Certified Alcoholism and Substance Abuse Counselor–Trainee (CASAC-T) in New York State and is currently pursuing a Master's in Applied Health Informatics. She also holds a B.S. in Psychology with a Minor in Business Administration and an M.S. in Cybersecurity, all from Fordham University.

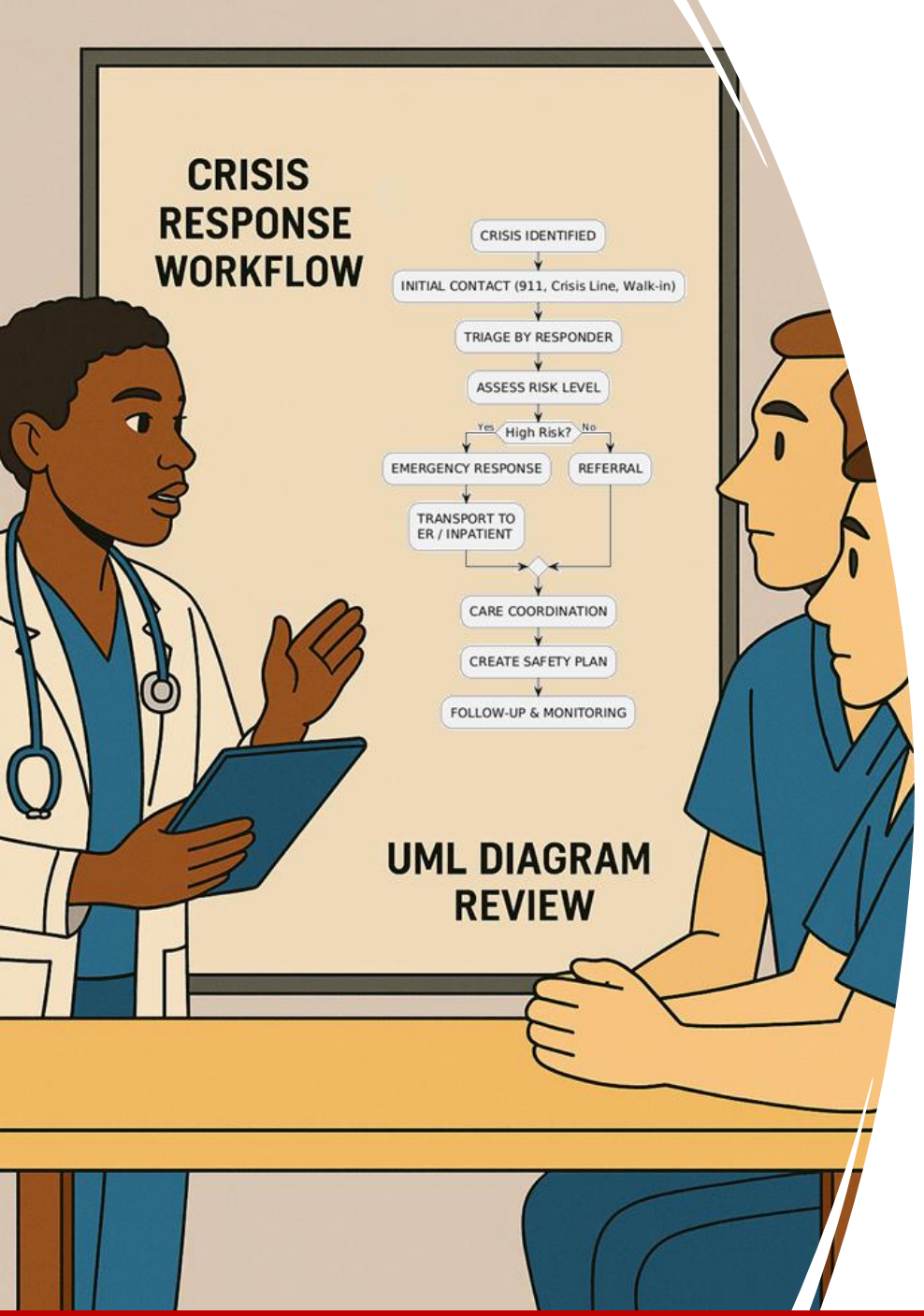
Her work focuses on using technology to improve care coordination and design systems that support mental health and recovery services.

## John Chelsom, Ph.D.

Program Director, Applied Health Informatics, Fordham University

John Chelsom has over 30 years of experience in health informatics. He holds a Ph.D. in Artificial Intelligence in Medicine from City University, London and a degree in Engineering Science from the University of Oxford. John has led the development of the first web-based Electronic Health Record (EHR) used in England's NHS, which became the foundation of the Summary Care Record.

He is the creator of the open source cityEHR system, used in hospitals and health informatics education. He is also an Adjunct Assistant Professor at the University of Victoria, Canada.




# Introduction

**Research Objective**: Evaluate AI tools for generating UML use case diagrams for Serious Mental Illness (SMI) crisis management.

**Purpose**: Accelerate development of care pathways for healthcare providers managing SMI crises.

**Focus**: Assess AI-generated UML diagrams for accuracy, clarity, efficiency, and cost.

**Approach**: Compare six AI-generated diagrams against a human-created benchmark.

# Background

**Serious Mental Illness (SMI) Crisis Management**:

- Requires detailed organization and procedures
- Needs clear care pathways for effective treatment
- Involves coordination between multiple healthcare providers

**Unified Modeling Language (UML)**:

- Standardized modeling language for system design
- Effective tool for visualizing software and system design
- Helps break down complex processes into manageable components

# Why UML for Healthcare?

- Creates visual representations of complex care relationships
- Simplifies the modeling of component relationships
- Shows relationships between users (providers) and the system of care
- Supports the development of clear care pathways
- Improves analysis by breaking down complex steps

# PlantUML


**What is PlantUML?**

An open-source, text-based UML diagramming tool supporting all standard UML types (class, use case, activity, sequence, etc.).

**Why use PlantUML?**

**Trusted in healthcare**: Proven reliable and syntactically correct in modeling complex healthcare systems.

**Agile-friendly**: Well-suited for incremental design and iterative system modeling using a simple, code-based approach.

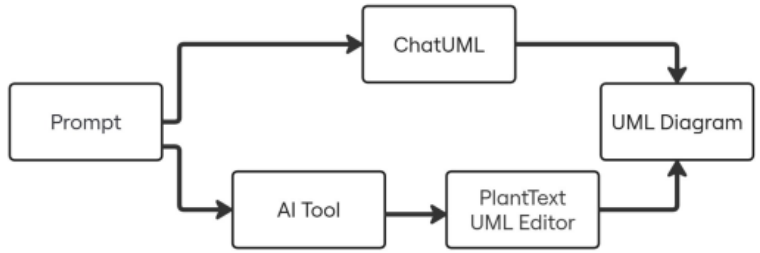
**AI-Compatible**: Works well with LLMs like ChatGPT and Claude for accurate, low-error UML generation.

**Widely adopted**: Preferred tool in AI-assisted UML modeling due to broad representation in training data.

**In Practice:** This study evaluated AI-generated UML diagrams to identify the tool that produced the most accurate, efficient and clearly structured representations of Serious Mental Illness (SMI) crisis workflows.

# UML Diagram Generation: Dual-Path Procedure

- **Methodology:** Prompts processed via two paths: directly through ChatUML or through AI Tool + PlantText UML Editor.
- **Testing Environment:** LG Gram laptop (Intel Evo i7, 16GB RAM) with stable internet connection.
- **Measurement:** Average response times from three repeated tests per method.
- **Benchmark:** Human-generated UML diagram created as quality reference standard.

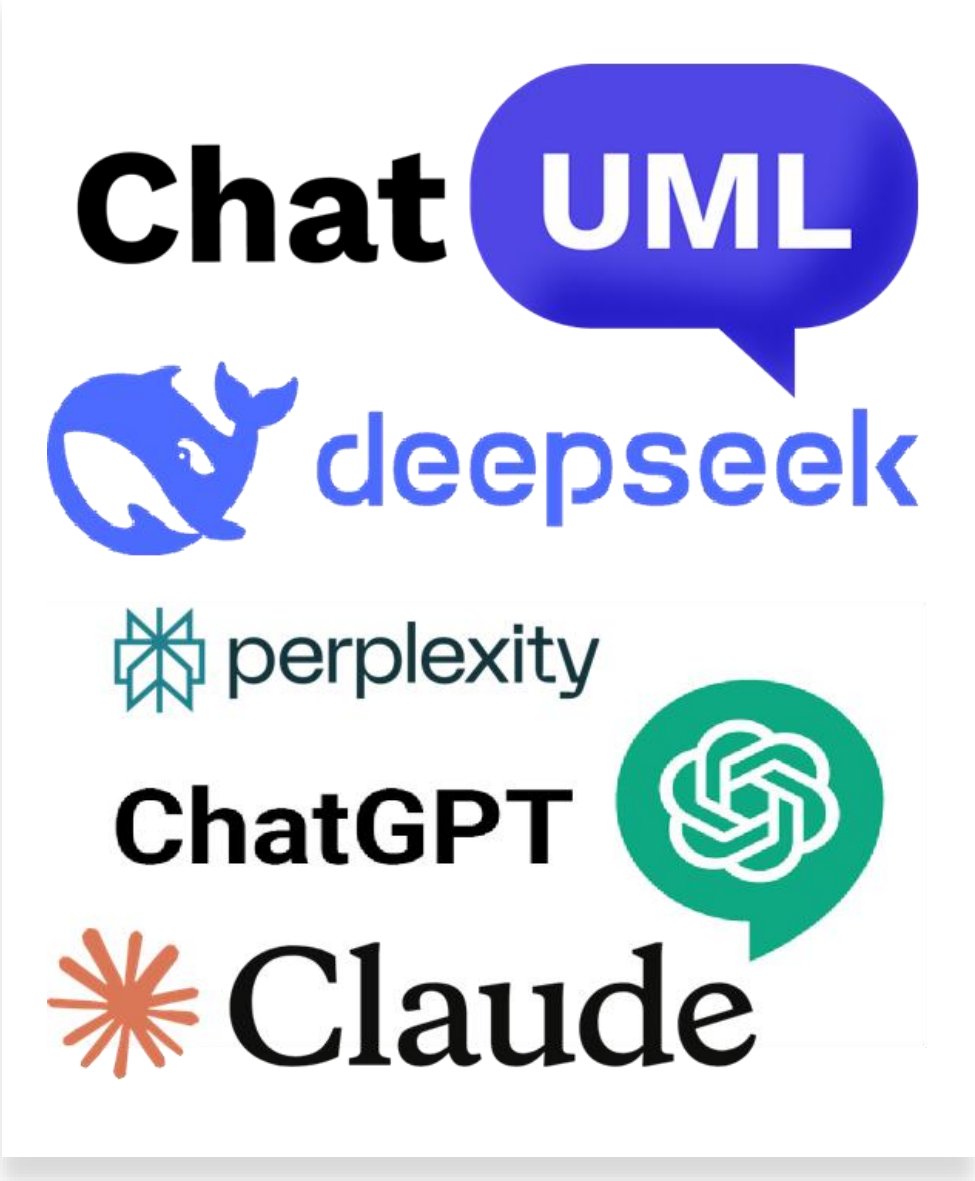


# AI Tools Evaluated:

1. ChatUML with Claude 3.5 Sonnet
2. ChatUML with ChatGPT 4o
3. ChatUML with DeepSeek V3
4. Claude 3.5 Sonnet (standalone)
5. ChatGPT 4o (standalone)
6. DeepSeek Reasoning with R1 via Perplexity
7. Human-created benchmark (for comparison)

*A consistent prompt related to SMI crisis management was used across all tools.*

**Data Source**: Medicaid Innovation Accelerator Program (IAP) list of behavioral health procedures.



# Evaluation Criteria

Each AI-Generated diagram was evaluated based on the below metrics, weighted according to their relevant importance.

**Technical Accuracy (40%)**:

- Adherence to relationships and prompt instructions
- Scale: 1-5 (5 = perfect alignment with requirements)

**Diagram Clarity (30%)**:

- Readability and usability of diagrams
- Scale: 1-5 (5 = well-organized with logical flow)

**Time Efficiency (20%)**:

- Total time from prompt initiation to final output
- Scale: 1-5 (5 = under 10 seconds)

**User Cost (10%)**:

- Expenses required for each tool
- Scale: 1-5 (5 = free)

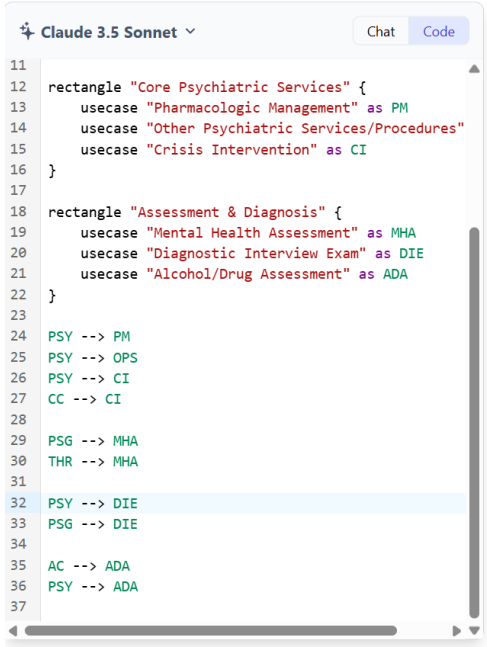
# Prompt Design

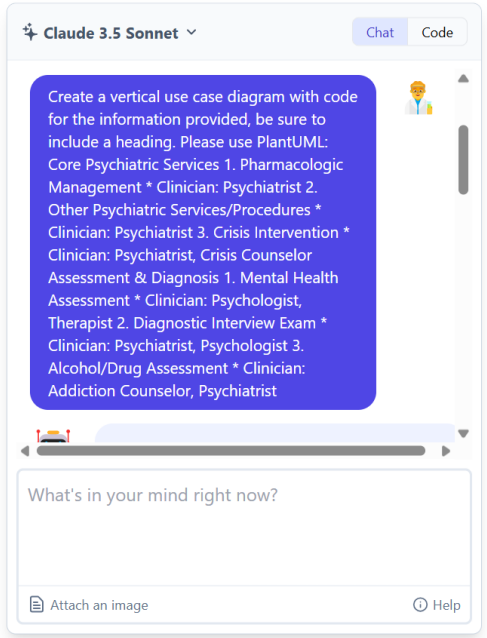
Original prompt was seven packages, truncated to two for presentation brevity.

**Truncated Prompt for the Practical Example (two packages)**:

Create a vertical use case diagram with code for the information provided, be sure to include a heading. Use PlantUML:

```
Core Psychiatric Services
1. Pharmacologic Management * Clinician: Psychiatrist
2. Other Psychiatric Services/Procedures * Clinician: Psychiatrist
3. Crisis Intervention * Clinician: Psychiatrist, Crisis Counselor
Assessment & Diagnosis
1. Mental Health Assessment * Clinician: Psychologist, Therapist
2. Diagnostic Interview Exam * Clinician: Psychiatrist, Psychologist
3. Alcohol/Drug Assessment * Clinician: Addiction Counselor, Psychiatrist
```




# Key Findings: Technical Accuracy

- **ChatGPT (in both environments) and native Claude Sonnet 3.5** did not consistently follow the prompt instructions.
- **ChatGPT** incorrectly generated horizontal diagrams despite being asked for vertical ones.
- **Native Claude** defaulted to its built-in Mermaid tool instead of using PlantUML as specified.

Technical Accuracy was the highest weighted variable having the most critical impact on care planning design and outcome.

# Key Findings: Diagram Clarity

- **ChatUML (Claude)** and **ChatUML (DeepSeek)** produced the clearest diagrams.
- **Native Claude's** diagrams were practically illegible due to distracting colors and small text.
- **DeepSeek (Perplexity)** and the **human-created diagram** also exhibited high diagram clarity.

This highlights the importance of formatting, structure and visual presentation in Diagram Clarity.

# Key Findings: Time Efficiency

- The study found that AI tools can significantly improve diagram creation efficiency.
- **ChatUML using Claude 3.5 Sonnet** and **DeepSeek Reasoning with R1 (via Perplexity)** performed best in this regard.
- The manual creation of the benchmark diagram was considerably more time-consuming.

The quicker the tool generates accurate diagrams, the more efficiently it can support iterative workflows and real time decision-making during development.

# Key Findings: User Cost

- Most expensive was **ChatGPT Native** at $20 per month.
- Middle tier included **Claude, ChatGPT,** and **DeepSeek within ChatUML**, which used a credit system (250 credits for $6.99), with each AI costing 3 credits per request at the time of the study.
- Free options were **Human-generated, Claude Native,** and **DeepSeek/Perplexity Native**.

The study concluded that the cost of AI tools varies significantly, with well performing free options available alongside more expensive subscription models. **ChatUML** offered a mid-range option with its credit-based system.

# Results: Top Performers

**ChatUML with Claude 3.5 Sonnet (4.9/5)**

Perfect technical accuracy (5/5)

Excellent diagram clarity (5/5)

Very efficient (5/5)

Moderate cost (4/5)

**DeepSeek Reasoning with R1 via Perplexity (4.8/5)**

Perfect technical accuracy (5/5)

Excellent diagram clarity (5/5)

Good efficiency (4/5)

Free to use (5/5)

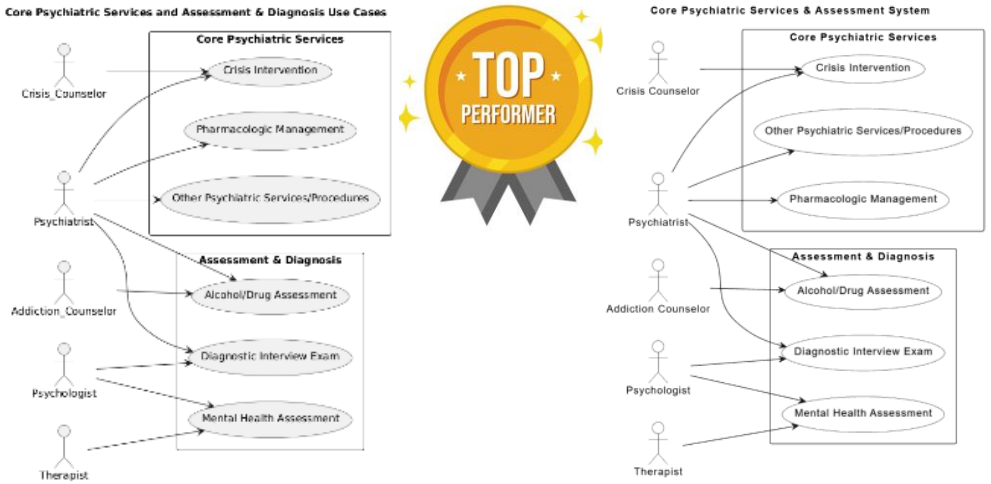
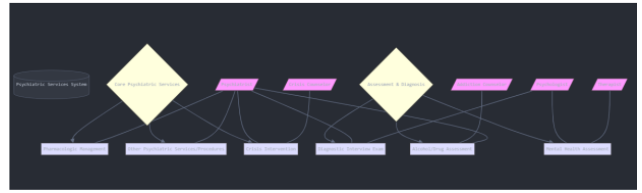



Table 1: UML Use Case Diagram Scorecard

| AI Tool | Technical Accuracy | Diagram Clarity | Time Efficiency | User Cost | Weighted |
|---|---|---|---|---|---|
| ChatUML, Claude | 5 | 5 | 5 | 4 | 4.9 |
| ChatUML, ChatGPT | 2 | 2 | 5 | 4 | 2.8 |
| ChatUML, DeepSeek | 3 | 5 | 2 | 4 | 3.5 |
| Human | 5 | 5 | 1 | 5 | 4.2 |
| Claude | 1 | 2 | 4 | 5 | 2.3 |
| ChatGPT | 3 | 4 | 4 | 1 | 3.3 |
| DeepSeek, Perplexity | 5 | 5 | 4 | 5 | 4.8 |

# Results: Low Performers

| |
|---|
| **ChatUML with ChatGPT 4o (Score: 2.8/5)** |
| Poor technical accuracy (2/5) |
| Poor diagram clarity (2/5) |
| Excellent efficiency (5/5) |
| Moderate cost (4/5) |
| **Claude 3.5 Sonnet Standalone (Score: 2.3/5)** |
| Very poor technical accuracy (1/5) |
| Poor diagram clarity (2/5) |
| Good efficiency (4/5) |
| No cost (5/5) |

# Conclusion: Need for HITL

- Limitations in AI-generated outputs demonstrate the critical need for the role of a human.
- There is a clear need for a human-in-the-loop (HITL) approach to ensure oversight, contextual judgment, and ethical alignment.
- This is particularly true in healthcare settings, where high-stakes decisions demand precision and should leave no room for error.
- HITL serves as a control for AI errors and a method of enhancing AI effectiveness in healthcare.

# Conclusion: Recommendations

**Top Recommendation**: ChatUML with Claude 3.5 Sonnet

- Best balance of accuracy, clarity, and efficiency
- Moderate cost is justified by performance

**Alternative Recommendation**: DeepSeek Reasoning with R1 via Perplexity

- Nearly equal performance to top choice
- Free to use, slightly slower due to extra step with PlantText

**Key Insight**: AI significantly reduces diagram creation time while maintaining quality, but human oversight remains essential for accuracy.

# Future Directions

- Incorporate more domain-specific knowledge into AI LLMs
- Develop AI capability to update diagrams based on real-time data
- Gather feedback on user interactions with AI-generated diagrams
- Further integrate HITL (Human-in-the-Loop) principles
- Expand to other healthcare documentation applications

***AI-assisted UML modeling shows significant promise for accelerating healthcare system design while maintaining human control through oversight.***

# References


[1] F. Röhricht, et al. "Implementation of a novel primary care pathway for patients with severe and enduring mental illness," *BJPsych Bulletin*, vol. 41, no. 6, pp. 314-319, Dec. 2017, doi: 10.1192/pb.bp.116.055830.

[2] Object Management Group (OMG), *Unified Modeling Language (OMG UML), Version 2.5.1*, 2017. [Online]. Available: https://www.omg.org/spec/UML/2.5.1/PDF retrieved: January, 2025

[3] H. Koç, A. M. Erdoğan, Y. Barjakly, and S. Peker, "UML diagrams in software engineering research: A systematic literature review," *Proceedings*, vol. 74, no. 1, p. 13, 2021. [Online]. Available: https://doi.org/10.3390/proceedings2021074013 retrieved: January, 2025

[4] J. Cámara, J. Troya, L. Burgueño, et al., "On the assessment of generative AI in modeling tasks: an experience report with ChatGPT and UML," *Softw. Syst. Model.*, vol. 22, pp. 781–793, 2023. [Online]. Available: https://doi.org/10.1007/s10270-023-01105-5 retrieved: January, 2025

[5] A. Conrady and J. Cabot, "From image to UML: First results of image-based UML diagram generation using LLMs," *arXiv:2404.11376 [cs.SE]*, 2024. [Online]. Available: https://doi.org/10.48550/arXiv.2404.11376 retrieved: January, 2025

[6] ChatUML, "ChatUML," Retrieved: Feb. 13, 2025. [Online]. Available: https://chatuml.com/ retrieved: January, 2025

[7] Centers for Medicare & Medicaid Services, "Using data analytics to better understand Medicaid populations with serious mental illness," *Medicaid Innovation Accelerator Program*, n.d. [Online]. Available: https://www.medicaid.gov/state-resource-center/innovation-accelerator-program/iap-downloads/program-areas/data-analytics-smi-tech-resource.pdf retrieved: January, 2025

[8] M. I. Ahmed, B. Spooner, J. Isherwood, M. Lane, E. Orrock, and A. Dennison, "A systematic review of the barriers to the implementation of artificial intelligence in healthcare," *Cureus*, vol. 15, no. 10, p. e46454, 2023. [Online]. Available: https://doi.org/10.7759/cureus.46454 retrieved: January, 2025

[9] E. Sezgin, "Artificial intelligence in healthcare: Complementing, not replacing, doctors and healthcare providers," *Digital Health*, vol. 9, p. 20552076231186520, 2023. [Online]. Available: https://doi.org/10.1177/20552076231186520 retrieved: January, 2025

[10] DeepSeek AI, "DeepSeek-V3 technical report," *arXiv*, Dec. 26, 2024. [Online]. Available: https://arxiv.org/abs/2412.19437 retrieved: January, 2025